CS 134: Networks	**Problem set 10**
Prof. Yaron Singer	Due: **Due Wednesday, 4/19/17 at noon**

The first two questions of this problem set are drawn from *Understanding Machine Learning: Theory and Algorithms.*

1. Suboptimality of k-means (Understanding Machine Learning, 22.8 Q1) (20 points) For every parameter $t > 1$, show that there exists an instance of the k-means problem for which the k-means algorithm (might) find a solution whose k-means objective is at least $t \cdot OPT$, where OPT is the minimum k-means objective.

Answer: Consider a set of 4 points that form a rectangle $(0,0), (0,h), (w,0), (w,h)$ were $w > h$. The optimal centroid solutions OPT, that could be found by k-means would be $(0, h/2), (w, h/2)$. However, k-means could also converge to the centroids of $(w/2, 0), (w/2, h)$ because $w > h$, the non-optimal solution $= t * OPT$, where t is a constant greater than 1. Because we can increase w arbitrarily, which keeps OPT the same while increasing the sub-optimal solution, we can grow t to be arbitrarily large. Therefore for any $t > 1$ we can create an existence/solution of the k-means problem whose k-means objective is at least $c \cdot OPT$

Solution courtesy of Nicholas Kochanek.

2. k-means Might Not Necessarily Converge to a Local Minimum (Understanding Machine Learning, 22.8 Q2) (20 points) Show that the k-means algorithm might converge to a point which is not a local minimum. Hint: Suppose that $k = 2$ and the sample points are $\{1, 2, 3, 4\} \subset \mathbb{R}$; suppose we initialize the k-means with the centers $\{2, 4\}$; and suppose we break ties in the definition of C_i by assigning i to be the smallest value in $\text{argmin}_j \mid\mid x - \mu_j \mid\mid$.

Solution by Matthew Huang

Suppose $k = 2$, and the points are $\{1, 2, 3, 4\}$. Initialize $\mu_1 = 2, \mu_2 = 4$. Breaking ties by choosing the smaller index i, we form clusters $C_1 = \{1, 2, 3\}, C_2 = \{4\}$. But in the next iteration, μ_1, μ_2 will not change, so no re-assignments happen and the k-means has “converged." However, this is not a local minima. Consider what happens when μ_1 is decreased slightly: $\mu_1 = 2 - \delta$. Then, at the next iteration, we will assign $C_1 = \{1, 2\}, C_2 = \{3, 4\}$, and will converge at this clustering with $\mu_1 = 1.5, \mu_2 = 3.5$.

For the first clustering, our loss is $1 + 1 = 2$, while for the second clustering, our loss is $4 \cdot 0.25 = 1$. Therefore, since a small perturbation to μ improved our loss, the k-means algorithm did not converge to a local minimum.

3. Ethics, revisited (30 points) During the ethics lecture, we discussed a number of different strategies that Facebook could use to suppress the spread of fake news through its network. Briefly explain one of these strategies in a single paragraph. Then, in two further paragraphs, explain whether or not you think Facebook is morally obligated to implement the strategy. Defend your answer, drawing on material from the ethics lecture. Your answer should be between 400 and 600 words.

4. Investigating Parameters of K-Means (30 points) **Solution by Samuel Cheng**

a. We find $|V| = 2992, |E| = 4046$.

b. Average shortest pairs distance: 11.5096.

c. Under min: 2, under max: 4, under mean: 3.

d. For the min and mean norm, the second cluster is closer, and under the max norm, the first cluster is closer.

e. The center of [2, 3, 4, 8, 20, 26] is 3.

f. *No answer required apart from code*

g. *Student answers can vary greatly, but common points should include:*

- *Clusters get smaller and more uniform as k increases.*
- *The objective function decreases as k increases.*
- *Clusters are much more uniform for the max norm, and become less so for the mean norm. The min norm almost always creates several huge clusters and many small ones.*
- *The objective function is very high for the min norm, is low for the max norm, and is lowest for the mean norm.*